

含有平面结构场景的捆绑调整

谢远帆¹ 吴毅红¹ 范力欣²

摘要 捆绑调整是计算机视觉中三维结构恢复过程的重要步骤. 捆绑调整通常将空间中点与点坐标的调整视为相互独立的过程, 但是在包含有自然物和人工物的场景中, 由于存在多余的自由度, 这种调整方法会导致优化结果偏离真值. 提出了一种带有共面约束和平面夹角约束的捆绑调整, 用于优化带有平面的场景. 借助新的参数化方法, 共面约束和夹角约束可以方便地进行表示, 并且带有这两类约束的捆绑调整求解过程, 仍然是一个无约束的非线性最小二乘问题. 实验结果表明, 这种带有先验信息的捆绑调整提供了对结构的更准确估计. 由于先验信息的加入, 增强型法方程的维度变高, 借助了稀疏的求解技术和预条件子方法, 大大降低了求解时间. 最后, 为了在实际应用中能够自动寻找出夹角约束, 提出了一种基于最大完全图的贪心方法, 该方法尽可能多地保留了夹角约束.

关键词 捆绑调整, 三维重建, L-M (Levenberg-Marquardt) 方法, 稀疏方程求解

引用格式 谢远帆, 吴毅红, 范力欣. 含有平面结构场景的捆绑调整. 自动化学报, 2014, 40(8): 1601–1611

DOI 10.3724/SP.J.1004.2014.01601

Bundle Adjustment for Scenes Containing Planes

XIE Yuan-Fan¹ WU Yi-Hong¹ FAN Li-Xin²

Abstract Bundle adjustment has been considered as one of the most important components in computer vision systems where three-dimensional structures are needed. A general bundle adjustment can optimize the coordinates of space points independently. But for a scene composed of both natural and structured objects, this often leads to an over parametrization the result and which deviates from truth. In this paper, a bundle adjustment with planar constraints and angle constraints is proposed for recovering the structures of environments with planes. By the aid of a new parametrization, the optimization remains an unconstrained non-linear least squares problem even if these two kinds of constraints are added. Experiments show that this new bundle adjustment method with prior knowledge provides an accurate estimation of the structures. Since prior information is added, the dimensionality of the augmented normal equation increases. A sparse solver is used after preconditioning in order to alleviate this problem. Moreover, a graph based angle constraint inference is devised for automatically finding constraints in a greedy manner once all planes are identified. This greedy method can preserve as many angle constraints as possible.

Key words Bundle adjustment, 3D reconstruction, Levenberg-Marquardt (L-M) method, sparse solver

Citation Xie Yuan-Fan, Wu Yi-Hong, Fan Li-Xin. Bundle adjustment for scenes containing planes. *Acta Automatica Sinica*, 2014, 40(8): 1601–1611

在三维重建的应用领域, 恢复带有自然物和人工物场景的结构是一个常见的问题. 处于具有先验结构表面上的空间点坐标由于满足等式约束, 因而自由度为 2. 但是在大部分从相机运动恢复结构 (Structure from motion, SFM) 方法中, 空间点被

抽象为三维向量, 并且在位置估计过程中相互独立. 这使得重建结果与真实三维结构和相机运动参数之间存在偏差. 因此, 设计一个包含有先验信息的捆绑调整是一项值得研究的问题. 许多工作已经探讨过在捆绑调整中通过添加约束的方式来引入先验^[1–11]. 一个直观的方法是在代价函数中引入加权的惩罚函数. 这种方法适用性广, 能解决许多形式的约束问题. 但是, 此类方法提供的结果往往不能严格满足所有约束, 并且惩罚函数的形式和权值都会影响最终结果.

如果从空间点坐标满足的等式约束出发, 可以通过求解方程和变量替换的方法去除原始变量之间的等式依赖关系, 进而为代价函数选择一组新的参数. 这种重新参数化的方法可以使约束得到严格满足, 并且最终仍然是求解一个无约束的非线性最小二乘问题.

收稿日期 2012-12-13 录用日期 2013-08-13
Manuscript received December 13, 2012; accepted August 13, 2013

国家重点基础研究发展计划 (973 计划) (2012CB316302), 国家自然科学基金 (61070107) 资助

Supported by National Basic Research Program of China (973 Program) (2012CB316302), and National Natural Science Foundation of China (61070107)

本文责任编辑 查红彬

Recommended by Associate Editor ZHA Hong-Bin

1. 中国科学院自动化研究所模式识别国家重点实验室 北京 100190, 中国 2. 诺基亚研究中心 坦佩雷 33720, 芬兰

1. National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
2. Nokia Research Center, Tampere 33720, Finland

这种严格满足约束的解,在真实感建模和摄影测量等应用中意义重大.虽然在一般形式的约束下,对代价函数进行重新参数化非常困难,但对于线性等式约束,可以很方便地实现重新参数化的过程.

本文提出了一种新的带有共面和夹角约束的捆绑调整方法.首先,根据空间点之间的共面约束,将共面点集参数化为所处平面的参数和所有点在该平面上的二维坐标参数两部分;然后,利用平面间夹角约束对相关平面的参数做进一步调整;最后,将新形式的空间点参数、平面参数和相机运动参数代入重投影误差之和这一代价函数中,得到含有先验信息的优化目标.代价函数极值的求解可采用 L-M (Levenberg-Marquardt) 方法,但由于平面参数的引入,增强型法方程的维度比无先验的方法高,因此我们采用稀疏的求解方法,以降低复杂度.

对比无约束的捆绑调整和其他融合先验信息的捆绑调整,本文提出的方法有以下三点优势: 1) 共面约束严格满足,夹角约束也得到了尽可能多的满足; 2) 共面点的参数化方法可以同时处理共面约束和平面间的夹角约束; 3) 优化问题的求解仍然是无约束的非线性最小二乘,迭代的每一步只需要求解稀疏的增强型法方程.

1 相关工作

在场景结构或相机运动已有先验信息的情况下,设计一个包含这些先验信息的捆绑调整是非常必要的.已有许多研究致力于利用先验信息来提高三维结构或相机运动参数估计的精度.这些方法可分为带有约束的运动和带有约束的三维结构两类.

第一类方法根据视觉系统的相机配置或从其他传感器获得的测量量对相机的运动参数进行约束^[1-4]. Lhuillier 提出了一种带相机位置约束的捆绑调整用于在重建过程中融合全球定位系统信息^[1].该方法在调整过程中,将相机位置约束在全球定位系统提供的位置上或其附近,通过约束相机位置参数和空间点位置两者更新量的模,使重投影误差的变化在每次迭代中有上限. Wong 等充分利用了转台重建方法对相机位置的约束^[2].通过将相机位置参数化在一个圆周上,带约束的捆绑调整提供了比无约束的捆绑调整更准确的结果. Di 等为火星探测器着陆地点建立三维地图任务开发了带有约束的捆绑调整^[3].该方法从立体视觉相机位置配置出发,导出了两个相机之间的几何约束,并将该约束用于参数化相机运动参数. Börlin 等比较了在捆绑调整中两种不同的旋转矩阵参数化对最终结果的影响^[4].

第二类方法主要讨论具有先验结构信息的捆绑调整^[5-11]. Zhou 等提出了通过空间平面诱导的单应进行内参数自标定和三维结构重建^[5].利用该方

法避免了平面引起的退化和去除运动前景对重建的影响,因此达到鲁棒效果.但其方法中并未讨论夹角的约束. Shan 等提出了用带有约束的捆绑调整来改进人脸建模^[6].在这个方法中,空间点被限制在一张参数化曲面上.在目标函数中引入线性化的惩罚函数后,用拉格朗日乘数法求解. Fua 提出了用带正规化项的捆绑调整对人头建模^[7].当曲面光滑时,该方法通过在目标函数中加入正规化项的方法,避免了曲面的过度变形.虽然该方法仅在人的头部数据上进行了验证,但可以应用在其他类似场合中. Szeliski 等将平面约束用在重建过程中^[8].该方法首先从平面诱导的单应导出基本矩阵的形式,其次平面约束也作为捆绑调整的约束加入到优化过程中,但是该方法的捆绑调整过程将相机位置估计和空间结构估计分为两个阶段交错进行,因此并非最优. Bartoli 等分析了共面约束并且对满足约束的几何元素提出了最小化的参数化方法^[9].由于平面结构的参数化在射影空间完成,因此该方法并未研究平面夹角的约束. Gerke 提出了一种带有水平、垂直和直角约束的捆绑调整^[10].该方法用于从倾斜角度拍摄的航拍图片进行三维重建.从这些约束出发,处于水平和垂直方向上线段的端点被重新参数化并用于进一步的优化,但在近景测量的情况下,水平和垂直约束一般不容易自动得到. McGlone 讨论了如何通过带约束的捆绑调整对特征匹配和几何关系假设进行验证,进而区分内点和外点^[11].

当需要从没有方位信息的图片恢复带有平面结构的场景时,融合有结构先验信息的捆绑调整将更具优势.对于度量重建问题,本文提出了一种可以同时表达共面约束和夹角约束的参数化方法.将此参数化后的空间点坐标代入重投影误差之和这一代价函数中,可以使平面参数、空间点参数和相机运动参数通过 L-M 迭代过程,同时得到优化.较之前的融合先验信息的方法,我们在共面约束基础上,进一步引入了夹角约束,增加了方法的精度和实用性.

2 针对包含平面结构的带约束的捆绑调整

无约束情况下捆绑调整问题定义如下^[12-13]:

$$\arg \min_{X_i, C_j} \left\{ \sum_i \sum_j \|\text{Proj}(X_i, C_j) - M_{ij}\|^2 \right\} \quad (1)$$

其中, X_i 是第 i 个空间点坐标, M_{ij} 是 X_i 在第 j 个相机 C_j 下的观测值.本节主要讨论如何利用先验信息将共面点进行参数化并代入式 (1) 中求解,以达到在优化过程中融合结构上的先验信息.第 2.1 节讨论如何参数化一个共面点集合与该集合确定的平面,并且讨论了初始化这些参数的方式.第 2.2 节讨论式 (1) 中的代价函数被新参数替换之后的形式以

及求解方法. 第 2.3 节讨论如何从空间点中寻找夹角约束.

2.1 参数化

为了能够同时处理共面约束和夹角约束, 一个共面点集所确定的平面 p 的参数定义为如下形式:

$$p = \{O, N_0, N_1\} = \{O, R(\gamma)e_0, R(\gamma)e_1\} \quad (2)$$

其中, O 是平面 p 上的点, N_0 和 N_1 是两个单位正交向量且两者的外积是该平面的法向量, $R(\gamma)$ 是旋转矩阵, e_0 和 e_1 是优化过程中保持不变的两个单位正交向量. 在本文方法中, 旋转矩阵采用指数映射表示方式, 定义为 $R(\gamma) = e^{[\gamma]_{\times}}$, 其中 γ 是该旋转矩阵对应的角轴表示, $[\gamma]_{\times}$ 是定义该旋转的 3×3 反对称阵. 指数形式与其他表示形式之间的关系参考文献 [14]. 由于 e_0 和 e_1 是常向量, 平面 p 的法向量 $(R(\gamma)e_0) \times (R(\gamma)e_1) = R(\gamma)(e_0 \times e_1)$ 则完全由 γ 决定. 由于 e_0 和 e_1 在给定 N_0 和 N_1 后由 $R(\gamma)$ 完全确定, 当选用另一个 γ 作为法向量参数时, 可以通过选择相应的 e_0 和 e_1 保持法向量不变. 对应于这种平面的参数化方式, 在 p 上的一点 X_i 具有如下形式:

$$X_i = O + u_0^i N_0 + u_1^i N_1 = O + [N_0 \ N_1] u_i \quad (3)$$

其中, $u_i = [u_0^i \ u_1^i]^T$ 是该点在该平面的二维坐标系下的坐标.

图 1 是平面的参数和其上一点的参数. 这种参数化方法为同时处理共面约束和夹角约束提供了简便的方法. 从式 (2) 和式 (3) 得到共面点集中的某点 X_i 可以表示为:

$$X_i(O, \gamma, e_0, e_1, u_i) = O + R(\gamma)[e_0 \ e_1] u_i \quad (4)$$

令 X_i^1 和 X_i^2 分别是参数化为 $\{X_i^1(O^1, \gamma^1, e_0^1, e_1^1, u_i^1)\}$ 和 $\{X_i^2(O^2, \gamma^2, e_0^2, e_1^2, u_i^2)\}$ 的两个共面点集. 由 X_i^1 和 X_i^2 确定的两平面的夹角为:

$$\theta = \langle R(\gamma^1)(e_0^1 \times e_1^1), R(\gamma^2)(e_0^2 \times e_1^2) \rangle \quad (5)$$

如果 γ^1 和 γ^2 是互相独立的变量, 则两个平面的夹角在优化过程中将发生变化. 假如两个平面的法向量由同一个初始化为 γ^1 的 γ 参数决定, 即两个共面点集被重新参数化为 $\{X_i^1(O^1, \gamma, e_0^1, e_1^1, u_i^1)\}$ 和 $\{X_i^2(O^2, \gamma, R'e_0^2, R'e_1^2, u_i^2)\}$, 其中 $R' = R^{-1}(\gamma^1)R(\gamma^2)$. 此时不论 γ 取何值, 两个平面的夹角仍然保持为式 (5). 即两个法向量相对运动关系不变, 优化过程中呈刚体形式在空间中运动. 这个例子解释了本文方法中如何参数化需要保持夹角不变的两个或多个共面点集. 将式 (4) 代入式 (1) 可以得到一个新的参数下具有共面约束和夹角约束的代

价函数. 通过最小化该代价函数, 无约束的空间点、平面上的空间点、平面参数和相机运动参数都可以得到优化.

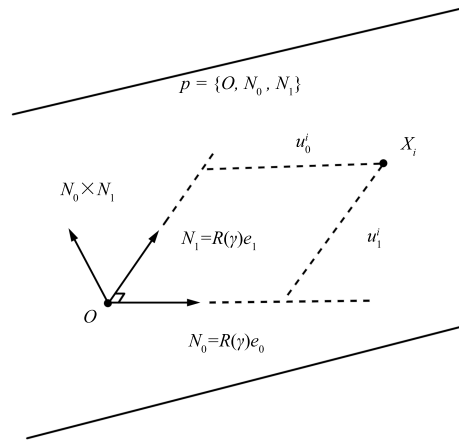


图 1 平面及其上的点的参数化方法

Fig. 1 Parameters for a plane and a space point on it

2.1.1 共面点集参数的初始化

在实际应用中, 如果有点集 $\mathcal{P} = \{X_i, i = 1, \dots, n\}$ 被认为是共面点集, 则平面的参数可以按如下方式得到:

首先, O 被初始化为点集的重心, 因此

$$O = \frac{1}{n} \sum_{i=1}^n X_i \quad (6)$$

式中, O 仅是一个参数, 而并不一定要在优化过程中保持是该平面的重心. 在得到 O 之后, 构建如下 \mathcal{P} 的数据矩阵:

$$M = [\tilde{X}_1 \ \dots \ \tilde{X}_n] = [X_1 - O \ \dots \ X_n - O] \quad (7)$$

设 M 的奇异值分解形式为:

$$M = U \Sigma V^T = [U_1 \ U_2 \ U_3] [\Lambda_{3 \times 3} | 0_{3 \times (n-3)}] V^T \quad (8)$$

N_0 和 N_1 可被初始化为 U 的前两列, 即:

$$N_0 = U_1, N_1 = U_2 \quad (9)$$

假设 $e_0 = [1 \ 0 \ 0]^T$, $e_1 = [0 \ 1 \ 0]^T$, 根据式 (2) 和式 (9), 不难发现 $R = U$.

当平面参数 p 已经从矩阵 \mathcal{P} 得到, 点 X_i 在该平面上的二维坐标 u_i 可由下式求得:

$$u_i = \begin{bmatrix} \langle N_0, X_i - O \rangle \\ \langle N_1, X_i - O \rangle \end{bmatrix} \quad (10)$$

如果由该点集确定的平面与其他点集确定的平面间没有夹角约束, 则该平面在本文中被称为自由平面.

2.1.2 相互间具有夹角约束的平面簇的参数化方法

一个平面簇在本文中定义为一组两两之间夹角固定的平面, 类似于第 2.1 节介绍的例子, 其中所有平面的法向量由同一个旋转参数 γ 以及各自的 e_0^i 和 e_1^i 决定. 因此, 一个平面簇的参数是所有该簇的平面的参数 O 和所有平面共有的一个 γ . 由于 O 的确定在这个情况下与第 2.1.1 节的方法相同, 所以本节主要讨论如何确定公共的旋转 $R(\gamma)$ 和每个平面的 e_0^i 和 e_1^i . 不失一般性, 我们以法向量线性无关的三个平面作为例子阐述参数化的方法. 这种方法很容易扩展到多于三个平面的情况.

假设 $p_i = \{O^i, N_0^i, N_1^i\}, i = 1, 2, 3$ 是由三个共面点集确定的三个空间平面. 三个平面两两之间的夹角在优化过程中需要保持不变.

如图 2 所示, 三个角度约束是:

$$\begin{cases} \langle n_1, n_2 \rangle \geq \cos \theta_{12} \\ \langle n_2, n_3 \rangle \geq \cos \theta_{23} \\ \langle n_1, n_3 \rangle \geq \cos \theta_{13} \end{cases} \quad (11)$$

其中, $n_i = N_0^i \times N_1^i = R^i(e_0^i \times e_1^i)$ 是三个平面的单位法向量. 为了保持两两平面间的夹角不变, 三个平面的法向量应当由同一个旋转 $R(\gamma)$ 参数化. 这个簇的旋转参数可以按如下方式得到:

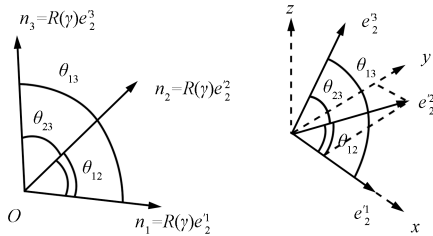


图 2 三个平面的法向量配置以及相应的 e_2^i

Fig. 2 The configuration of the normals of the three planes and the corresponding e_2^i 's

假设 $e_2^i, i = 1, 2, 3$ 是三个单位向量, 使得 Re^{i_2} 是 p_i 的法向量. 将 $Re^{i_2}, i = 1, 2, 3$ 代入式 (11), 得:

$$\begin{cases} \langle e_2^1, e_2^2 \rangle \geq \cos \theta_{12} \\ \langle e_2^2, e_2^3 \rangle \geq \cos \theta_{23} \\ \langle e_2^1, e_2^3 \rangle \geq \cos \theta_{13} \end{cases} \quad (12)$$

在确定满足式 (12) 的三个向量 e_2^1, e_2^2 和 e_2^3 后, 则:

$$R = [n_1 \ n_2 \ n_3][e_2^1 \ e_2^2 \ e_2^3]^{-1} \quad (13)$$

如图 2 所示, e_2^1 和 e_2^2 可以分别选为 $[1 \ 0 \ 0]^T$ 和 $[\cos \theta_{12} \ \sin \theta_{12} \ 0]^T$, 而 e_2^3 可以由 $\|e_2^3\| = 1, (n_1 \times n_2)^T n_3 = (e_2^1 \times e_2^2)^T e_2^3$ 和式 (12) 确定.

对于平面个数大于 3 的情况, 首先构建矩阵 D :

$$D = [n_1 \cdots n_N] \quad (14)$$

使用如下的方法可以很容易确定含有 N 个平面的簇的旋转参数:

1) 如果 $\text{rank}(D) = 1$, 即所有平面具有相同的法向量, 此时寻找到这个簇的公共旋转矩阵与单个平面的参数化是一样的;

2) 如果 $\text{rank}(D) = 2, n_1$ 和 n_2 可以选为 D 中两个不同的向量, 而 n_3 可以选为 $n_1 \times n_2$, 因此 e_2^3 应为 $e_2^1 \times e_2^2$.

3) 如果 $\text{rank}(D) = 3, n_1, n_2$ 和 n_3 可以选为 D 中线性无关的三个向量.

4) 将簇中所有平面初始的 R^i 用新 R 替换时, 不应改变平面 p_i 的法向量, 因此将 $e_j^i = R^{-1}R^i e_j^i, j = 0, 1$ 选为各个平面的新的 e_j^i . 显然, $R(e_0^i \times e_1^i) = N_0^i \times N_1^i$.

属于这个簇的所有平面都具有 $p_i = \{O^i, R(\gamma)e_0^i, R(\gamma)e_1^i\}$ 形式的参数. 由于所有平面具有相同的旋转参数 γ , 因此在优化过程中可以很精确地满足夹角约束.

在实际应用中, 在点集估计出来的平面法向量之间, 往往并不严格满足从先验知识中得到的那些约束. 所以需要将实际数据中得到的受噪声影响的一个平面簇法向量与先验地设定的法向量通过一个旋转变换尽量对齐. 对齐的标准是对应的法向量之间夹角最小. 因此, 求得每个 e_2^i 对应的 n_i 后, 公共的旋转 R 可以通过求解下列问题得到:

$$\arg \min_{\gamma} \sum_i -\langle R(\gamma)e_2^i, n_i \rangle \quad (15)$$

得到 $R(\gamma)$ 后, 平面法向量 n_i 将被 $R(\gamma)e_2^i$ 取代. 同时, 对应的 e_0^i 和 e_1^i 也需要重新计算, 使 $e_0^i \times e_1^i = e_2^i$.

至此, 我们得到了用于后续捆绑调整的满足共面约束和夹角约束的三维结构参数集合. 对比无约束的捆绑调整和其他融合先验信息的捆绑调整, 本文方法有以下三点优势: 1) 共面约束严格满足, 并且夹角约束也得到了尽可能多的满足; 2) 共面约束和平面间的夹角约束在同一种参数化形式下得到了处理; 3) 将具有新参数形式的空间点坐标代入代价函数后, 优化问题的求解仍然是无约束的非线性最小二乘.

2.2 被重新参数化后的代价函数以及求解方法

当所有共面点集按照共面约束和夹角约束参数化成式 (4) 后, 式 (1) 中被新参数替换后的代价函数可用无约束的方法求解. 假设空间结构由 M 个无约

束的三维点, N 个自由平面和 K 个平面簇组成, 记 X_i 是第 i 个无约束点的坐标, Y_i^j 是第 j 个自由平面上的第 i 个点, Z_i^{jk} 是第 k 个平面簇中的第 j 个平面上的第 i 个点. 在新的参数下, 代价函数为:

$$\arg \min_P \sum_{k=1}^K \sum_j \sum_i \sum_l \|\text{Proj}(Z_i^{jk}, C_l) - M_{il}^{jk}\|^2 + \sum_{j=1}^N \sum_i \sum_l \|\text{Proj}(Y_i^j, C_l) - M_{il}^j\|^2 + \sum_{i=1}^M \sum_l \|\text{Proj}(X_i, C_l) - M_{il}\|^2 \quad (16)$$

其中, M_{il} 、 M_{il}^j 和 M_{il}^{jk} 分别是相应空间点在相应图像下匹配的特征点坐标. 由于式 (16) 中的参数 P 具有分块结构, $P = [p_c \ p_q \ p_\gamma \ p_p \ p_v \ p_u \ p_x]$, 其中 p_c 是所有相机参数的集合, p_q 是所有属于某个簇的平面的 O 参数组成的向量, p_γ 是由 K 个簇的 K 个公共旋转参数组成的向量, p_p 是由 N 个自由平面的参数组成的向量, p_v 是由所有属于某个簇的点的 u 坐标 (如式 (3) 所示) 组成的向量, p_u 是由所有属于某个自由平面的点的 u 坐标组成的向量, p_x 是所有自由点的坐标组成的向量. 式 (16) 中的问题可以通过常用的数值方法求解. 本文采用 L-M 方法, 迭代的每一步需要求解如下增强型法方程:

$$(J^T J + \lambda I) \delta = -J^T E \quad (17)$$

其中, J 是所有空间点在相机下的投影对参数向量的导数. 在新的参数下, $J^T J$ 是分块稀疏矩阵, 其上三角部分如下.

$$\begin{bmatrix} C & C_Q & C_\Gamma & C_P & C_V & C_U & C_X \\ & Q & Q_\Gamma & & Q_V & & \\ & & \Gamma & & \Gamma_V & & \\ & & & P & & P_U & \\ & & & & V & & \\ & & & & & U & \\ & & & & & & X \end{bmatrix} \quad (18)$$

由于 V 、 U 和 X 都是块对角可逆矩阵, 块矩阵 $\text{diag}\{V, U, X\}$ 的 Schur 补的上三角部分为:

$$A = \begin{bmatrix} \tilde{C} & \tilde{C}_Q & \tilde{C}_\Gamma & \tilde{C}_P \\ & \tilde{Q} & \tilde{Q}_\Gamma & \\ & & \tilde{\Gamma} & \\ & & & \tilde{P} \end{bmatrix} \quad (19)$$

其中,

$$\tilde{C} = C - C_V V^{-1} C_V^T - C_U U^{-1} C_U^T - C_X X^{-1} C_X^T \quad (20)$$

$$\tilde{C}_Q = C_Q - C_V V^{-1} Q_V^T \quad (21)$$

$$\tilde{C}_\Gamma = C_\Gamma - C_V V^{-1} \Gamma_V^T \quad (22)$$

$$\tilde{C}_P = C_P - C_U U^{-1} P_U^T \quad (23)$$

$$\tilde{Q} = Q - Q_V V^{-1} Q_V^T \quad (24)$$

$$\tilde{Q}_\Gamma = Q_\Gamma - Q_V V^{-1} \Gamma_V^T \quad (25)$$

$$\tilde{\Gamma} = \Gamma - \Gamma_V V^{-1} \Gamma_V^T \quad (26)$$

$$\tilde{P} = P - P_U U^{-1} P_U^T \quad (27)$$

如果没有任何先验约束加入, 式 (19) 将变成

$$A = C - C_X X^{-1} C_X^T \quad (28)$$

这就是传统无约束方法中需求解的约简相机参数系统^[12].

假设重建的场景结构和相机参数包括了 L 个具有 n 个参数的相机, 属于 K 个簇的平面共 P 个, 自由平面共 N 个. 在这种情况下, 式 (19) 中的 A 是 $nL + 3P + 3K + 6N$ 阶实对称阵. 随着越来越多的平面加入到优化中, Schur 补的维度将随着平面个数线性增长. 因此, 迭代每一步求解法方程的代价要比无约束的相应求解代价高. 为了降低求解复杂度, 我们使用了稀疏的求解技术^[15] 和基于图分割的预条件子方法^[16]. 稀疏求解技术使矩阵 A 的对称分解复杂度仅与对称分解因子的非零元个数有关; 而基于图分割的预条件子技术对求解变元进行重新排序, 进一步降低了对称分解的复杂度.

2.3 从点集中寻找夹角约束

共面点集可以通过平面模型拟合或手工方法选择出来. 当所有的共面点集都找到后, 每个点集都按照自由平面方法进行参数化. 通过测试所有平面两两间法向量的夹角是否与先验的夹角之一接近, 可以建立一个图 $G = [V \ E]$. 在这个图中, 每一个顶点代表一个平面, 每一条边表示两顶点所代表的平面之间有夹角约束, 而一个完全子图表明属于该子图的顶点两两之间有夹角约束, 因此构成一个平面簇. 由于要找出满足所有约束的一种参数化方法非常困难^[9], 因此我们采取了一种贪心的方法来保留尽可能多的夹角约束. 首先, 在这个图中找到最大的完全子图, 并且将属于该完全子图的顶点对应的所有平面作为一个平面簇进行参数化; 然后, 这个完全子图将被移除出原有的图 G . 用贪心的方法进行迭代直到剩下的顶点全部为孤立顶点; 最后, 所有的孤立顶点对应的平面都作为自由平面进行参数化. 图 3 为

一个简单的例子. $\{p_3, p_4, p_5\}$ 和 $\{p_6, p_7, p_8\}$ 分别是两个长方体的三个面, $\{p_1, p_2\}$ 是两个任意放置的平面. 该例子中先验的角度是 0 和 $\pi/2$. 指定的 e_2^i 是 $[1\ 0\ 0]^T$, $[0\ 1\ 0]^T$ 和 $[0\ 0\ 1]^T$. 对应于该配置的约束图中的虚线边对应的夹角约束没有考虑在内.

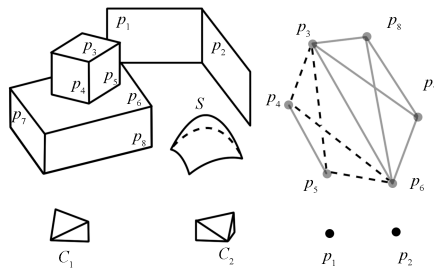


图 3 具有先验信息的空间结构和对应的约束图

Fig. 3 A structure with a prior information and its corresponding constraint graph

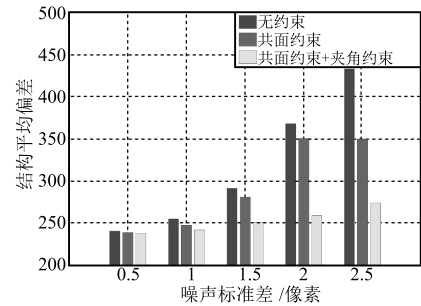
3 实验结果

包含结构信息的优化方法的有效性在模拟数据和实际应用中都得到了验证. 无约束的捆绑调整 (BA_n)、仅平面约束的捆绑调整 (BA_p)、平面和夹角约束的捆绑调整 (BA_{pa}) 三种方法在实验中进行比较. 三种方法的比较在具有真值的模拟数据上进行. 评价方法优劣的标准是重建结果和真实值之间的差. 第 3.2 节是针对标定块的实验, 进一步比较三种方法的性能, 并且通过建筑物的重建的例子, 展示本文方法的实用性.

3.1 模拟数据实验

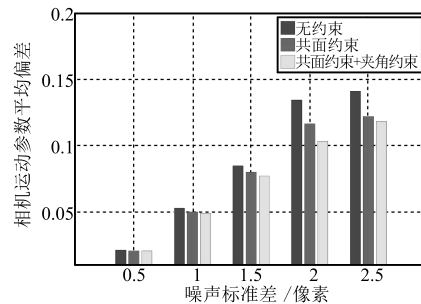
模拟数据通过 SFM 方式得到. 首先, 根据空间点位置和相机运动参数, 生成各相机下的合成图像; 然后, 对图像上空间点的投影位置加上均值为 0 的高斯噪声的扰动. 高斯噪声的标准差为 0.5~2.5, 以 0.5 为间隔, 从 0.5 变化到 2.5. 根据图像上对应点的投影坐标以及内参数, 可以对模拟的场景进行三维重建. 经过三种方法优化后, 各个方法给出的空间点位置和相机运动参数估计值与真实值进行比较, 说明约束对重建结果的影响. 估计值与真实值的平均偏差用估计参数和真实参数之间的欧氏距离的平均值衡量. 为了去除度量重建之间相似变换的影响, 第一个相机的运动参数设置为 $[R\ t] = [I\ 0]$, 并且第一个与第二个相机之间的基线长度归一化为 1. 图 4 为本次实验的结果. 从图 4 可以看出, 带约束的优化结果比不带约束的优化结果离真值更近. 不带约束的优化结果随着噪声水平的增加也增加得更快. 随着越来越多的约束加入到优化中, 估计值和真值之间的偏差越来越小, 并且随噪声水平的增长速度也

变得更加缓慢, 其原因在于随着约束的增多, 场景结构的已知成分就越大.



(a) 空间结构的平均偏差

(a) Average deviation of structure



(b) 运动参数的平均偏差

(b) Average deviation of motion

图 4 估计值与真值间的平均偏差

Fig. 4 Average deviation of estimated parameters from truth

为了进一步说明本文方法的特性, 我们比较了三种方法在不同视数、不同最大基线长度和不同共面点比例下, 对同一场景的三维结构和相机运动参数进行优化的结果. 实验中所使用的模拟数据由准确的三维结构、相机运动参数和空间点在相机上的投影加上 0 均值高斯噪声得到. 每组数据具有不同的视数、最长重建基线和共面点占有所有空间点的比例. 每个配置下生成 10 次数据用于优化, 取偏差的平均值作为测试方法的估计值偏差. 在讨论重建参数准确度和某个因素的关系时, 将该因素相同的所有测试结果进行平均. 例如, 当需要考察重建准确度与最大基线长度的关系时, 是将所有具有最大基线长度的结果进行平均. 这样处理更能说明这种关系在各种情况下的一致性. 图 5 说明了随着视数、最大基线长度和共面点比例的增加, 三种方法优化后空间结构估计的准确度均会增加. 并且约束越充分, 结构估计的准确度越高. 这也是由于先验信息的加入, 使得部分结构已知的结果. 图 6 表明了视数、最大基线长度和共面点比例增加的情况下, 相机运动参数估计的准确度也有所增加. 这是由于结构的改进, 对相机运动参数的估计有积极作用.

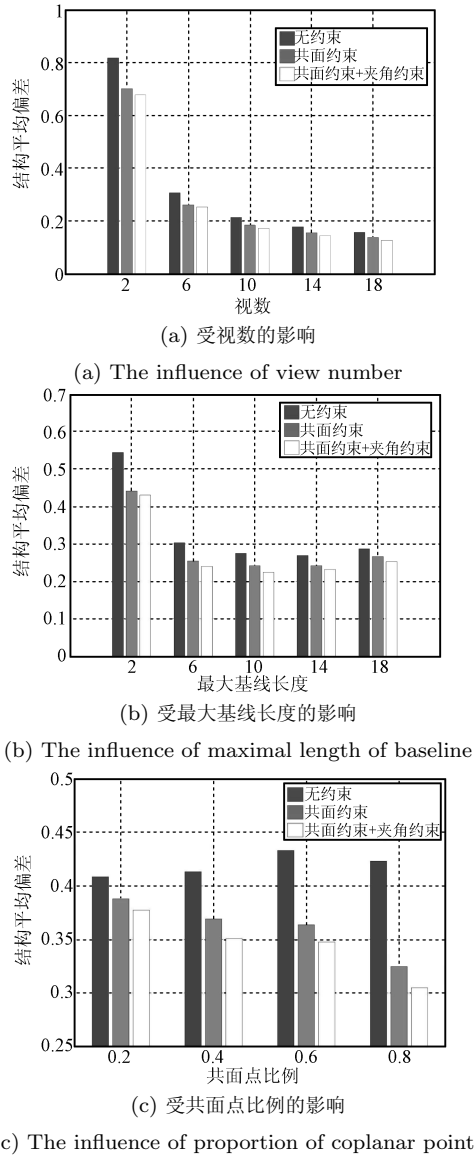


图 5 空间结构的平均偏差

Fig. 5 Average deviation of structure

图 7 显示了经过重新参数化后的空间结构与真值之间的差. 值得注意的是, 图 7(c) 反映了随着共面点比例的升高, 初值也和真值更加接近. 因此说明迭代次数的减少和迭代初值的改进密切相关. 从图 8 可以看出, 带约束方法比无约束方法收敛的平均迭代次数少.

为了评估稀疏方法的时间开销, 测试有结构约束方法的时间复杂度, 用不同视数和空间平面个数的模拟数据. 在测试数据中, 式 (19) 中的矩阵 A 的阶数是 $7m \times 6n$, 其中, m 是视数, n 是平面个数. 由于在实际重建应用中, 矩阵 A 往往比较稀疏, 因此在高斯消元或对称分解过程中可以仅针对非零元进行操作. 本文采用对称分解对方程进行求解. 稠密的对称分解方法所需时间与矩阵维度的三次方

成正比, 而稀疏的分解方法与分解结果的非零元个数成正比, 因此大幅度降低了时间复杂度. 另外, 对矩阵行列进行重新排列可以对分解过程进行并行化处理, 从而降低分解时间. 但是问题的性质从根本上决定了稀疏性带来的性能提升, 因此不同问题中稀疏解法的加速比也会相差很大.

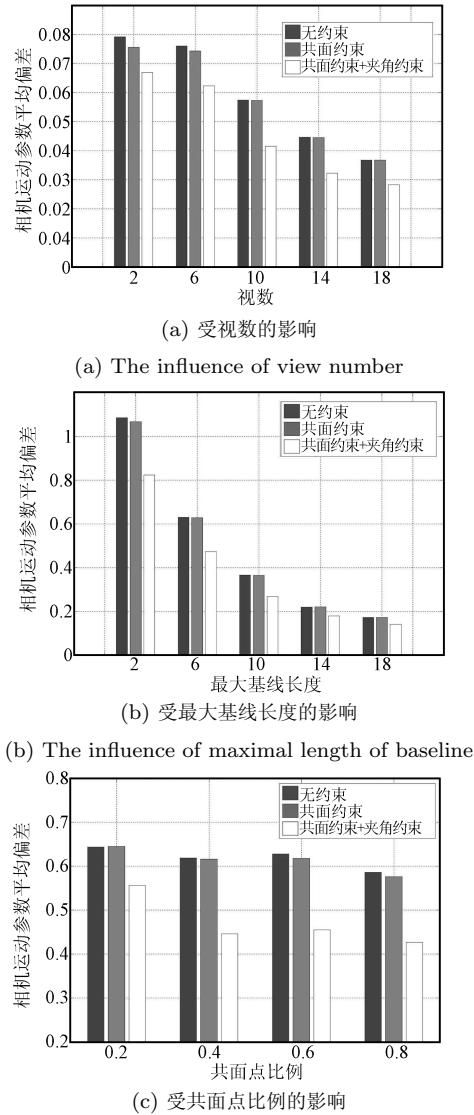
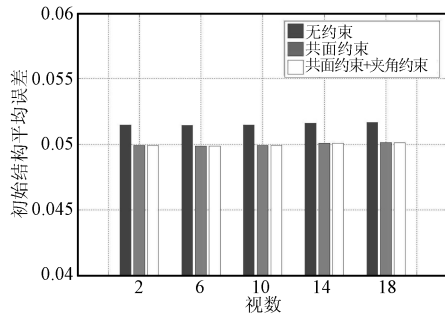


图 6 运动参数的平均偏差

Fig. 6 Average deviation of camera motion

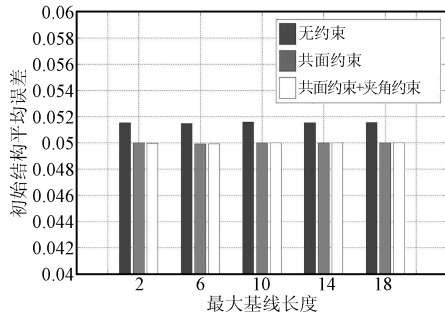
为此, 我们对稀疏性质不同的两个问题进行了求解时间上的比较. 问题 1 是一个法方程较为稀疏的重建问题, 每个空间点的观测数约为总视数的 $1/10$. 问题 2 是一个法方程全是非零元的重建问题, 所有的空间点在所有图片下有匹配的特征点. 这两个问题的求解时间见表 1 和表 2, 括号内数据是使用稠密求解方法的时间. 从表中可以看出, 稠密方法的求解时间仅与矩阵维度有关, 与问题的稀疏性无关.

使用稀疏的求解方法可以使求解时间大幅降低, 并且减缓所需时间随 A 阶数的增加而增加的速度。



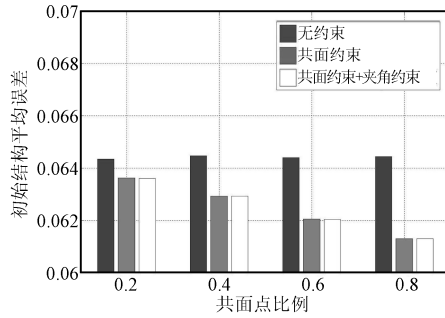
(a) 受视数的影响

(a) The influence of view number



(b) 受最大基线长度的影响

(b) The influence of maximal length of baseline



(c) 受共面点比例的影响

(c) The influence of proportion of coplanar points

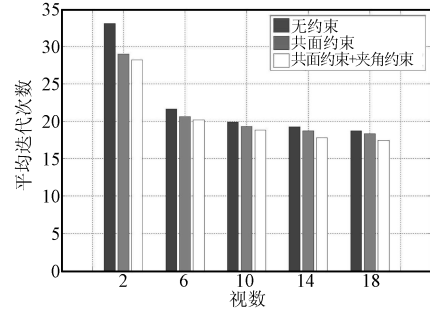
图 7 初始空间结构的平均偏差

Fig. 7 Average deviation of initial structure

3.2 实际应用测试

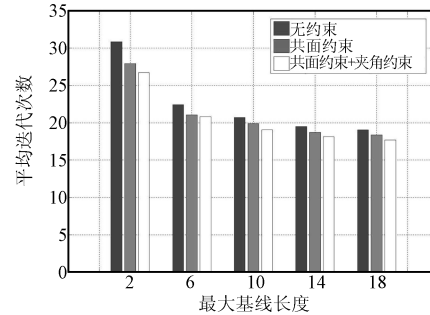
3.2.1 在标定块上的测试

本节利用恢复标定块的三维结构测试, 对这三种方法进行了比较. 如图 9 所示, 标定块上的点分为四类. 对于无约束的方法 BA_n , 所有点均认为是无约束的空间点. 对于仅具有平面约束的方法 BA_p , 三个正方形框上以及内部的点被认为是三个共面点集. 对于具有平面以及平面夹角约束的方法 BA_{pa} , 正方形框上以及内部的点被认为是三个共面点集, 并且平面的法向量两两垂直. 在后两种方法中, 框外的点被认为是自由点集合.



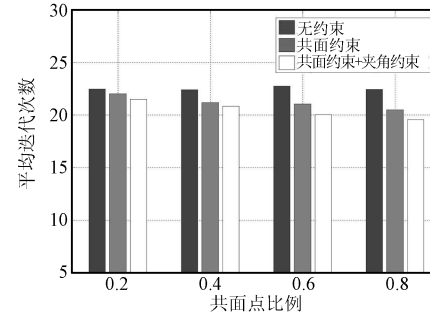
(a) 受视数的影响

(a) The influence of view number



(b) 受最大基线长度的影响

(b) The influence of maximal length of baseline



(c) 受共面点比例的影响

(c) The influence of proportion of coplanar points

图 8 平均迭代次数

Fig. 8 Average steps of iteration before convergency

三个正方形框的边被用于衡量各方法对结构恢复的准确度. 为了消除相似变换对结果的影响, 评价中使用了图 9 中所示的线段长度比值. 这些比值的真实值是 1. 用不同方法重建得到的结果进行该比值的测量, 如表 3 所示. 从测量的均值和标准差可以看出, 在使用了包含先验信息的优化后, 重建准确度和精度都得到了提升.

3.2.2 在建筑物重建中的应用

由于许多建筑物表面是由平面拼接而成, 所以本文方法可以用来得到一个与该建筑物先验信息相吻合的三维结构. 在得到初始重建后, 我们通过两种方法得到点集的共面信息. 一种是利用 J-linkage 实

表 1 稀疏和稠密方法用于求解问题 1 的法方程时间 (ms)

Table 1 Time costs of augmented normal equation solving based on sparse and dense methods for Problem 1 (ms)

视数 \ 平面数	0	2	4	6	8	10	12	14	16	18	20
20	0.5 (0.8)	0.5 (1.1)	0.6 (1.4)	0.7 (1.8)	0.7 (2.2)	0.7 (2.7)	0.8 (3.2)	0.8 (3.8)	0.8 (4.5)	0.8 (5.0)	0.9 (5.7)
60	2.2 (20)	2.4 (22)	2.4 (24)	2.4 (26)	2.5 (29)	2.5 (31)	2.6 (34)	2.6 (37)	2.7 (40)	2.8 (42)	3.0 (45)
100	6.1 (89)	6.6 (95)	6.9 (108)	6.5 (112)	6.7 (118)	6.1 (122)	6.7 (132)	6.8 (136)	6.8 (143)	6.9 (149)	6.9 (162)

表 2 稀疏和稠密方法用于求解问题 2 的法方程时间 (ms)

Table 2 Time costs of augmented normal equation solving based on sparse and dense methods for Problem 2 (ms)

视数 \ 平面数	0	2	4	6	8	10	12	14	16	18	20
20	1.0 (0.8)	1.2 (1.1)	1.5 (1.4)	1.8 (1.7)	1.8 (2.2)	2.1 (2.7)	2.3 (3.2)	2.7 (3.8)	3.0 (4.5)	3.4 (5.3)	3.9 (6.2)
60	8.5 (20)	9.2 (22)	10 (24)	11 (26)	12 (29)	12 (31)	14 (34)	14 (37)	16 (40)	17 (43)	17 (46)
100	30 (91)	31 (96)	33 (100)	35 (106)	37 (112)	38 (120)	40 (126)	42 (132)	44 (138)	46 (146)	47 (155)

表 3 线段长度比测量值

Table 3 Length ratio of the segments measured

方法	s2/s1	s3/s1	s4/s1	s6/s5	s7/s5	s8/s5	s10/s9	s11/s9	s12/s9	均值	标准差
BA_n	0.995450	1.02070	1.01465	0.975152	0.977288	0.994831	0.979676	0.974925	1.00458	0.993029	0.0174973
BA_p	0.997801	1.01539	1.01206	0.983739	0.984721	0.997248	0.991344	0.986924	1.00629	0.997281	0.0117884
BA_{pa}	0.999215	1.01134	1.00870	0.989130	0.991947	0.998167	0.995963	0.991761	1.00341	0.998854	0.00771090

现共面点集的聚类^[17]. 但考虑到自动方法在实际应用中的适用性, 我们保留了手工方法, 即通过在图片上指定共面点集投影区域来产生点集共面信息. 得到初步的共面点集后, 进一步利用极大似然估计采样一致方法^[18] 将所选的潜在共面点集内点选出, 进行平面拟合. 当所有平面都初始化完成, 根据第 2.3 节所述夹角约束方法将会自动获得. 在优化结束时, 每个平面都可以生成三角网格, 用于纹理映射.

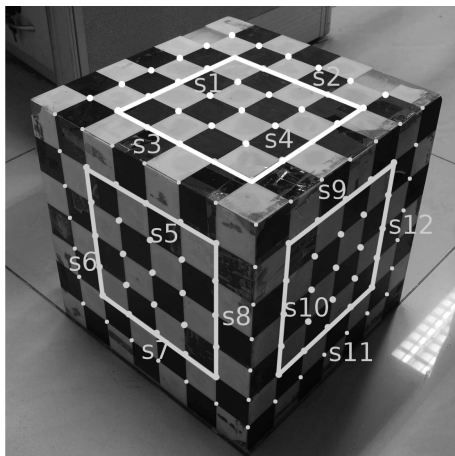


图 9 用于重建的点和测量的线段

Fig. 9 Points and segments to be measured

图 10(a) 中的教学楼具有 5 个不同朝向的外立面. 设定的先验夹角是 0 、 $\pi/4$ 和 $\pi/2$. 三维结构包含了 102 个视数、1 个约 8000 个点的平面簇和约 9000 个自由点, 优化时间为 295 秒. 图 10(a) 中的商场具有 12 个平面. 先验的角度是 0 和 $\pi/2$, 设定的 e_2^2 是 $[1 \ 0 \ 0]^T$ 和 $[0 \ 1 \ 0]^T$. 三维结构包含了 113 个视数, 9 个自由平面 (约 6000 点), 1 个包含有三个平面共约 3000 点的平面簇和大约 1 万个自由点, 优化时间为 305 秒. 在实际应用中可以发现, 平面间的夹角在矫正之前与先验的夹角之间有数度的偏差. 本文方法可以在严格满足共面约束和夹角约束条件下, 使相机运动参数和空间结构得到重投影误差最小的估计.

4 结论

通过对共面点进行满足共面约束和平面间夹角约束的重新参数化, 新的以重投影误差最小为目标的代价函数可以在严格满足两种约束的条件下, 对相机运动参数和包括自由空间点、共面的空间点以及空间平面在内的空间结构进行优化. 模拟实验和实际应用都验证了该方法的有效性. 该方法在城市环境等具有许多平面结构的场景恢复中, 有重要的应用价值. 这种方法也可以拓展到具有显式表达的

参数表面重建过程中。



(a) 用于重建的图像

(a) Images used for reconstruction



(b) 人机交互恢复出具有纹理的空间平面

(b) Textured planes recovered based on human machine interaction



(c) 自动平面模型拟合的平面恢复结果

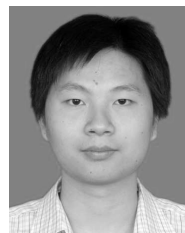
(c) Textured planes recovered based on automatic plane fitting

图 10 应用于建筑物重建时得到的表面模型

Fig. 10 Surface model acquired when applied to building reconstruction

References

- Lhuillier M. Fusion of GPS and structure-from-motion using constrained bundle adjustments. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Colorado Springs, USA: IEEE, 2011. 3025–3032
- Wong K H, Chang M M Y. 3D model reconstruction by constrained bundle adjustment. In: Proceedings of the 17th International Conference on Pattern Recognition. Cambridge, UK: IEEE, 2004, 3: 902–905
- Di K, Xu F, Li R. Constrained bundle adjustment of panoramic stereo images for Mars landing site mapping. In: Proceedings of the 4th International Symposium on Mobile Mapping Technology. Kunming, China: MMT, 2004. 29–31
- Börlin N, Grussenmeyer P, Eriksson J, Lindstrom P. Pros and cons of constrained and unconstrained formulation of the bundle adjustment problem. In: International Archives of ISPRS, 2004, XXXV(B3): 589–594
- Zhou Z H, Jin H L, Ma Y. Robust plane-based structure from motion. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Rhode Island, USA: IEEE, 2012. 1482–1489
- Shan Y, Liu Z C, Zhang Z Y. Model-based bundle adjustment with application to face modeling. In: Proceedings of the 8th IEEE International Conference on Computer Vision. Vancouver, Canada: IEEE, 2001, 2: 644–651
- Fua P. Regularized bundle-adjustment to model heads from image sequences without calibration data. *International Journal of Computer Vision*, 2000, 38(2): 153–171
- Szeliski R, Torr P H S. Geometrically constrained structure from motion: points on planes. *Lecture Notes in Computer Science*, 1998, 1506: 171–186
- Bartoli A, Sturm P. Constrained structure and motion from multiple uncalibrated views of a piecewise planar scene. *International Journal of Computer Vision*, 2003, 52(1): 45–64
- Gerke M. Using horizontal and vertical building structure to constrain indirect sensor orientation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2011, 66(3): 307–316
- McGlone J C. Bundle adjustment with geometric constraints for hypothesis evaluation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 1996. B3-III529–534
- Hartley R I, Zisserman A. *Multiple View Geometry in Computer Vision*. Cambridge: Cambridge University Press, 2004
- Triggs B, McLauchlan P F, Hartley R I, Fitzgibbon A W. Bundle adjustment—a modern synthesis. In: Proceedings of the International Workshop on Vision Algorithms: Theory and Practice. London, UK: IEEE, 2000. 298–372
- Wedderburn J H M. *Lectures on Matrices*. Providence: American Mathematical Society, 1934
- Timothy A D. *Direct Methods for Sparse Linear Systems*. Philadelphia: Society for Industrial and Applied Mathematics, 2006
- Rotkin V, Toledo S. The design and implementation of a new out-of-core sparse cholesky factorization method. *ACM Transactions on Mathematical Software*, 2004, 30(1): 19–46
- Toldo R, Fusiello A. Robust multiple structures estimation with J-linkage. In: Proceedings of the 10th European Conference on Computer Vision: Part I. Marseille. France: Springer-Verlag, 2008. 537–547
- Torr P H S, Zisserman A. MLESAC: a new robust estimator with application to estimating image geometry. *Computer Vision and Image Understanding*, 2000, 78(1): 138–156



谢远帆 中国科学院自动化研究所博士研究生。2007 年获中南大学信息科学与工程学院自动化专业学士学位。主要研究方向为基于视觉的同步定位与地图创建。本文通信作者。

E-mail: yfxie@nlpr.ia.ac.cn

(XIE Yuan-Fan Ph. D. candidate at the Institute of Automation, Chinese Academy of Sciences. He received his bachelor degree from Central South University in 2007. His research interest covers vision-based simultaneous localization and mapping. Corresponding author of this paper.)

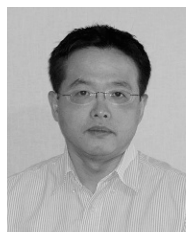


吴毅红 中国科学院自动化研究所研究员. 2001 年获中国科学院系统所理学博士学位. 主要研究方向为摄像机标定, 摄像机定位, 三维重建.

E-mail: yhwu@nlpr.ia.ac.cn

(**WU Yi-Hong** Professor at the Institute of Automation, Chinese Academy of Sciences. She received her

Ph.D. degree from the Institute of Systems Science, Chinese Academy of Sciences in 2001. Her research interest covers camera calibration, camera pose determination, and 3D reconstruction.)



范力欣 芬兰坦佩雷诺基亚研究中心媒体实验室首席科学家. 1998 年和 2002 年获新加坡国立大学硕士和博士学位. 主要研究方向为计算机视觉和模式识别.

E-mail: Lixin.fan@nokia.com

(**FAN Li-Xin** Principal research scientist in the Media Lab, Nokia Research Center, Tampere, Finland. He

received his master and Ph.D. degrees from National University of Singapore in 1998 and 2002. His research interest covers computer vision and pattern recognition.)